The Average Amount of Information Lost in Multiplication

Nicholas Pippenger
njp@princeton.edu

Department of Computer Science
Princeton University
35 Olden Street
Princeton, NJ 08540 USA

Abstract: We show that if X and Y are integers independently and uniformly distributed in the set $\{1, \ldots, N\}$, then the information lost in forming their product (which is given by the equivocation $H(X, Y \mid X \cdot Y)$), is $\Theta(\log\log N)$. We also prove two extremal results regarding cases in which X and Y are not necessarily independently or uniformly distributed. First, we note that the information lost in multiplication can of course be 0. We show that the condition $H(X, Y \mid X \cdot Y) = 0$ implies $2\log_2 N - H(X, Y) = \Omega(\log\log N)$. Furthermore, if X and Y are independent and uniformly distributed on disjoint sets of primes, it is possible to have $H(X, Y \mid X \cdot Y) = 0$ with $\log_2 N - H(X)$ and $\log_2 N - H(Y)$ each $O(\log\log N)$. Second, we show that however X and Y are distributed, $H(X, Y \mid X \cdot Y) = O(\log N / \log\log N)$. Furthermore, there are distributions (in which X and Y are independent and uniformly distributed over sets of numbers having only small and distinct prime factors) for which we have $H(X, Y \mid X \cdot Y) = \Omega(\log N / \log\log N)$.

1. Introduction

Let X and Y be random integers. We regard a multiplier as a deterministic channel whose input is the pair (X, Y) and whose output is the product $X \cdot Y$. The information lost in multiplication is, according to Shannon [S3], the equivocation $H(X, Y \mid X \cdot Y)$. From the definition of conditional entropy, we have

$$\begin{aligned} H(X, Y \mid X \cdot Y) &= H(X, Y, X \cdot Y) - H(X \cdot Y) \\ &= H(X, Y) - H(X \cdot Y), \qquad (1.1) \end{aligned}$$

where we have used the fact that the channel is deterministic ($X \cdot Y$ is determined by X and Y, so that $H(X, Y, X \cdot Y) = H(X, Y)$).

We first consider the case in which X and Y are independent and uniformly distributed on the set $\{1, \ldots, N\}$, so that $H(X, Y) = 2 \log_2 N$. We shall show in Section 2 that in this case we have

$$H(X, Y \mid X \cdot Y) = \Theta(\log \log N). \qquad (1.2)$$

If X and Y have arbitrary (that is, not necessarily independent or uniform) distributions on $\{1, \ldots, N\}$, then it is of course possible that $H(X, Y \mid X \cdot Y) = 0$. We may then ask how close $H(X, Y)$ can come to its maximum $2 \log_2 N$, while still achieving $H(X, Y \mid X \cdot Y) = 0$. We shall show in Section 3 that $H(X, Y \mid X \cdot Y) = 0$ implies that

$$2 \log_2 N - H(X, Y) = \Omega(\log \log N). \qquad (1.3)$$

Furthermore, by taking X and Y to be independent, with distributions concentrated on disjoints sets of primes, it is possible to achieve $H(X, Y \mid X \cdot Y) = 0$ with $\log_2 N - H(X)$ and $\log_2 N - H(Y)$ each $O(\log \log N)$, so that (1.3) is the best possible bound.

We shall also consider the distributions of X and Y that maximize the information loss. We shall show in Section 4 that for any distributions of X and Y on $\{1, \ldots, N\}$ we have

$$H(X, Y \mid X \cdot Y) = O(\log N / \log \log N). \qquad (1.4)$$

Furthermore, by taking X and Y to be independent, with distributions concentrated on integers having only small and distinct prime factors, we can achieve

$$H(X, Y \mid X \cdot Y) = \Omega(\log N / \log \log N),$$

so that (1.4) is the best possible bound.

Results concerning information flow through a multiplier have been used by Abelson and Andreae [A] and by Brent and Kung [B] to obtain lower bounds involving the area and time required for multiplication. Furthermore, the results in Section 4 give a lower bound to the number of ancillary lines required by a reversible multiplier (see Fredkin and Toffoli [F] for a discussion of reversible computation). This lower bound is achievable if multiplication is performed by a single gate; it is an open question whether it can be achieved if the multiplier is implemented using standard reversible gates, such as those proposed by Fredkin and Toffoli.

The proofs in this paper draw upon a variety of results from number theory. Many of these in turn rely on the prime-number theorem (first proved by Hadamard [H1] and independently by de la Vallée Poussin [V]) and its extension to primes in arithmetic progressions (first proved by de la Vallée Poussin [V]). While these deep theorems now have elementary proofs (due to Selberg [S1, S2] and Erdős [E1]), none of our results actually depend on theorems of this depth, and thus we shall take care to point out the simplest results that support our proofs.

2. The Uniform Distribution

Our goal in this section is to establish (1.2). For X and Y independent with the uniform distribution, we have

$$H(X, Y) = 2 \log_2 N.$$

Thus from (1.1) we have

$$H(X, Y \mid X \cdot Y) = 2 \log_2 N - H(X \cdot Y). \tag{2.1}$$

Define $m(N)$ by

$$m(N) = \#\{x \cdot y : 1 \le x \le N, 1 \le y \le N\}.$$

We have

$$H(X \cdot Y) \le \log_2 m(N).$$

Thus the bound

$$H(X, Y \mid X \cdot Y) = \Omega(\log \log N) \tag{2.2}$$

is a consequence of (2.1) and the following result.

Proposition 2.1: For any $\varepsilon > 0$, we have

$$m(N) \le \frac{N^2}{(\log N)^{\alpha - \varepsilon}} \tag{2.3}$$

for all sufficiently large N, where $\alpha = 1 - \log_2(e \ln 2) = 0.08607\ldots$.

This result is due to Erdős [E2], who also proved the matching bound

$$m(N) \geq \frac{N^2}{(\log N)^{\alpha+\varepsilon}}.$$

For completeness, we shall give a simple proof of this proposition.

Proof of Proposition 2.1: Let $f(n)$ denote the number of distinct prime factors in the integer $n \geq 1$. Let $\tau_k(x)$ denote the number of integers n in the interval $1 \leq n \leq x$ such that $f(n) = k$. Hardy and Ramanujan [H2] (Lemma B) have shown that there are absolute constants L and D such that

$$\tau_k(x) \leq \frac{Lx}{\ln x} \frac{(\ln\ln x + D)^{k-1}}{(k-1)!} \tag{2.4}$$

for all $k \geq 1$ and $x \geq 2$. Apart from an elementary precursor

$$\tau_1(x) = O\left(\frac{x}{\log x}\right)$$

to the prime-number theorem due to Chebyshev [C], their result relies only on the elementary estimates

$$\sum_{p \leq x} \frac{\ln p}{p} = \ln x + O(1)$$

and

$$\sum_{p \leq x} \frac{1}{p} = O(\log\log x)$$

(in which the sums are over primes p) due to Mertens [M]. We observe that (2.4) implies

$$\tau_k(x) \leq \frac{Mx}{\ln x} \frac{(\ln\ln x)^{k-1}}{(k-1)!} \tag{2.5}$$

for all $x \geq 2$ and $1 \leq k \leq 2\log_2 \ln x$, where $M = L\exp(2D\log_2 e)$.

Fix $0 < \delta < 1/6$. Define $m_1(N)$, $m_2(N)$ and $m_3(N)$ by

$$m_1(N) = \#\big\{(x, y) : 1 \leq x \leq N, 1 \leq y \leq N \text{ and } f(x) + f(y) \leq (1+2\delta)\log_2 \ln N\big\},$$

$$m_2(N) = \#\big\{z : 1 \leq z \leq n^2 \text{ and } f(z) \geq (1+\delta)\log_2 \ln N\big\}$$

and

$$m_3(N) = \#\big\{z : 1 \leq z \leq N^2 \text{ and } w^2 \mid z \text{ for some } w \text{ with } f(w) \geq \delta\log_2 \ln N\big\}.$$

Then we have

$$m(N) \leq m_1(N) + m_2(N) + m_3(N).$$

For if $z = x \cdot y$ is not counted by $m_1(N)$, then we have $f(x) + f(y) > (1 + 2\delta) \log_2 \ln N$. If in addition z is not counted by $m_2(N)$, then we have $f(z) < (1 + \delta) \log_2 \ln N$, and thus

$$f\bigl(\gcd(x, y)\bigr) > \delta \log_2 \ln N,$$

where $\gcd(x, y)$ denotes the greatest common divisor of x and y. If we now let w be the product of the distinct primes dividing $\gcd(x, y)$, then we have $w^2 \mid z$ and $f(w) \geq \delta \log_2 \ln N$, so that $z = x \cdot y$ is counted by $m_3(N)$. Thus it will suffice to show that $m_1(N)$, $m_2(N)$ and $m_3(N)$ each satisfy a bound of the form of that in (2.3).

For $m_1(N)$ we have

$$\begin{aligned} m_1(N) &\leq \sum_{i+j \leq (1+2\delta) \log_2 \ln N} \tau_i(N)\, \tau_j(N) \\ &\leq \frac{M^2\, N^2}{(\ln N)^2} \sum_{i+j=k \leq (1+2\delta) \log_2 \ln N} \frac{(\ln \ln N)^{k-2}}{(i-1)!\, (j-1)!} \\ &\leq \frac{M^2\, N^2}{(\ln N)^2} \sum_{i+j=k \leq (1+2\delta) \log_2 \ln N} \binom{k-2}{i-1} \frac{(\ln \ln N)^{k-2}}{(k-2)!} \\ &\leq \frac{M^2\, N^2}{(\ln N)^2} \sum_{i+j=k \leq (1+2\delta) \log_2 \ln N} \frac{(2 \ln \ln N)^{k-2}}{(k-2)!} \\ &\leq \frac{M^2\, N^2}{(\ln N)^2} \sum_{i+j=k \leq (1+2\delta) \log_2 \ln N} \left(\frac{2e \ln \ln N}{k-2}\right)^{k-2}, \end{aligned} \tag{2.6}$$

where we have used the definition of $m_1(N)$, the bound (2.5), the identity $a!/b!\,(a-b)! = \binom{a}{b}$, the inequality $\binom{a}{b} \leq 2^a$ and the inequality $a! \geq a^a/e^a$.

The summand in (2.6) increases with k for $k - 2 \leq 2 \ln \ln N$, and decreases thereafter. Since $k - 2 < (1 + 2\delta) \log_2 \ln N \leq 2e \ln \ln N$, the largest terms of the sum are those with the largest k. There are at most $\bigl((1 + 2\delta) \log_2 \ln N\bigr)^2 \leq (2 \ln \ln N)^2$ terms in all, and each term is at most

$$(2e \ln 2)^{(1+2\delta) \log_2 \ln N} = (\ln N)^{(1+2\delta)(2-\alpha)}.$$

Thus we obtain the bound

$$m_1(N) \leq \frac{M^2\, (2 \ln \ln N)^2\, (\ln N)^{2\delta(2-\alpha)}\, N^2}{(\ln N)^\alpha},$$

which is of the form desired, since if $2\delta(2 - \alpha) < \varepsilon$, the factors

$$M^2\, (2 \ln \ln N)^2\, (\ln N)^{2\delta(2-\alpha)}$$

in the numerator can be absorbed by the factor $(\ln N)^\varepsilon$ in the denominator of (2.3).

For $m_2(N)$ we have

$$\begin{aligned} m_2(N) &\leq \sum_{k \geq (1+\delta)\log_2 \ln N} \tau_k(N^2) \\ &\leq \frac{M\,N^2}{\ln N} \sum_{k \geq (1+\delta)\log_2 \ln N} \frac{(\ln\ln N)^{k-1}}{(k-1)!} \\ &\leq \frac{M\,N^2}{\ln N} \sum_{k \geq (1+\delta)\log_2 \ln N} \left(\frac{e\ln\ln N}{k-1}\right)^{k-1}, \end{aligned} \tag{2.7}$$

where we have used the definition of $m_2(N)$, the bound (2.5) the inequality $a! \geq a^a/e^a$.

The summand in (2.7) increases with k for $k-1 \leq \ln\ln N$, and decreases thereafter. Since $k-1 \geq (1+\delta)\log_2 \ln N - 1 \geq \ln\ln N$, the largest terms of the sum are those with the smallest k. There are at most $2e\log_2 \ln N$ terms with $k-1 < 2e\log_2 \ln N$, and each such term is at most

$$(e\ln 2)^{(1+\delta)\log_2 \ln N} = (\ln N)^{(1+\delta)(1-\alpha)}. \tag{2.8}$$

Furthermore, all the terms with $k-1 \geq 2e\log_2 \ln N$ are bounded by the terms of a geometric progression with ratio 1/2, and thus their sum is bounded by (2.8). Thus we obtain the bound

$$m_2(N) \leq \frac{M\,(1+2e\log_2 \ln N)\,(\ln N)^{\delta(1-\alpha)}\,N^2}{(\ln N)^{\alpha}},$$

which is of the form desired, since if $\delta(1-\alpha) < \varepsilon$, the factors

$$M\,(1+2e\log_2 \ln N)\,(\ln N)^{\delta(1-\alpha)}$$

in the numerator can be absorbed by the factor $(\ln N)^{\varepsilon}$ in the denominator of (2.3).

Finally, for $m_3(N)$ we have

$$\begin{aligned} m_3(N) &\leq \sum_{f(w) \geq \delta\log_2 \ln N} \frac{N^2}{w^2} \\ &\leq \sum_{w \geq w_0} \frac{N^2}{w^2} \\ &\leq \frac{N^2}{w_0}, \end{aligned} \tag{2.9}$$

where w_0 denotes the smallest integer w such that $f(w) \geq \delta\log_2 \ln N$. Clearly $w_0 = p_1 \cdots p_k$ is the product of the first $k = \lceil \delta\log_2 \ln N \rceil$ primes. If N is sufficiently large that there are fewer than $k/2$ primes that are less than $2^{2/\delta}$, then w_0 contains at least $k/2$ prime factors that are each at least $2^{2/\delta}$, and thus $w_0 \geq \ln N$. The bound (2.9) is therefore also of the desired form. This completes the proof of the proposition. □

Next we turn to establishing the upper bound

$$H(X, Y \mid X \cdot Y) = O(\log \log N). \tag{2.10}$$

To do this we use the formula

$$H(X, Y \mid X \cdot Y) = \sum_{1 \le x \le N} \sum_{1 \le y \le N} \Pr[X = x, Y = y]\, H(X, Y \mid X \cdot Y = x \cdot y). \tag{2.11}$$

Using the bound

$$\begin{aligned} H(X, Y \mid X \cdot Y = x \cdot y) &\le \log_2 \#\big\{(v, w) : 1 \le v \le N, 1 \le w \le N \text{ and } v \cdot w = x \cdot y\big\} \\ &\le \log_2 d(x \cdot y) \end{aligned}$$

(where $d(n)$ denotes the number of divisors of the integer n), we obtain

$$H(X, Y \mid X \cdot Y) \le \sum_{1 \le x \le N} \sum_{1 \le y \le N} \Pr[X = x, Y = y]\; \log_2 d(x \cdot y). \tag{2.12}$$

For X and Y independent with the uniform distribution, (2.12) becomes

$$H(X, Y \mid X \cdot Y) \le \frac{1}{N^2} \sum_{1 \le x \le N} \sum_{1 \le y \le N} \log_2 d(x \cdot y). \tag{2.13}$$

Since $\log_2 a$ is a concave function of a, the average of the logarithm in (2.13) is at most the logarithm of the average, and we obtain

$$H(X, Y \mid X \cdot Y) \le \log_2 \left(\frac{1}{N^2} \sum_{1 \le x \le N} \sum_{1 \le y \le N} d(x \cdot y) \right).$$

Since $d(x \cdot y) \le d(x) \cdot d(y)$, we obtain

$$\begin{aligned} H(X, Y \mid X \cdot Y) &\le \log_2 \left(\frac{1}{N^2} \sum_{1 \le x \le N} \sum_{1 \le y \le N} d(x) \cdot d(y) \right) \\ &= 2 \log_2 \left(\frac{1}{N} \sum_{1 \le n \le N} d(n) \right). \end{aligned} \tag{2.14}$$

We now use the asymptotic formula

$$\sum_{1 \le n \le N} d(n) = N \ln N + O(N)$$

due to Dirichlet [D2] (which is established simply by estimating the number of lattice points in the region bounded by the x-axis, the y-axis and the hyperbola $x \cdot y = N$). Substituting this result in (2.14) completes the proof of (2.10), which together with (2.2) establishes (1.2).

3. Multiplication without Loss of Information

Our goal in this section is to determine the maximum entropy that X and Y can have when $H(X, Y \mid X \cdot Y) = 0$. Let

$$\mathcal{W} = \{(x, y) : \Pr[X = x, Y = y] > 0\}$$

denote the support of the distribution of (X, Y), and let

$$\mathcal{M} = \{x \cdot y : 1 \leq x \leq N, 1 \leq y \leq N\}$$

be the range of the multiplication map $\mu : \{1, \ldots, N\} \times \{1, \ldots, N\} \to \{1, \ldots, N^2\}$ defined by $\mu(x, y) = x \cdot y$. Then $H(X, Y \mid X \cdot Y) = 0$ implies that μ restricted to $\mathcal{W}$ is injective, so that $\#(\mathcal{W}) \leq \#\mathcal{M} = m(N)$ and $H(X, Y) \leq \log_2 m(N)$. Proposition 2.1 thus shows that $H(X, Y \mid X \cdot Y) = 0$ implies (1.3).

To show that this result is the best possible, we let X and Y be independent and uniformly distributed over $\mathcal{X}$ and $\mathcal{Y}$, respectively, where $\mathcal{X}$ and $\mathcal{Y}$ are the sets of primes that are at most N and congruent to 1 and 3, respectively, modulo 4. To show that $\log_2 N - H(X)$ and $\log_2 N - H(Y)$ are each $O(\log\log N)$, it will suffice to show that $\#\mathcal{X} = \pi_{1,4}(N)$ and $\#\mathcal{Y} = \pi_{3,4}(N)$ are each $\Omega(N/\log N)$. This of course follows from the extention of the prime-number theorem to arithmetic progressions, but we can obtain what we need from the following simple result due to Shapiro [S4] (which is an elementary quantitative version of the theorem of Dirichlet [D1] on primes in arithmetic progressions). Let a and b be fixed with $\gcd(a, b) = 1$. Then

$$\sum_{\substack{p \leq x \\ p \equiv a \pmod{b}}} \frac{\ln p}{p} = \frac{\ln x}{\phi(b)} + O(1), \tag{3.1}$$

where $\phi(b)$ denotes Euler's totient function: the number of a in the range $0 < a < b$ such $\gcd(a, b) = 1$. To show that (3.1) implies

$$\pi_{a,b}(x) = \Omega\left(\frac{x}{\log x}\right), \tag{3.2}$$

we observe that (3.1) implies that

$$\sum_{\substack{x/A < p \leq x \\ p \equiv a \pmod{b}}} \frac{\ln p}{p} \geq \frac{\ln A}{\phi(b)} - 2B \tag{3.3}$$

for all $A > 1$, where B is a bound on the magnitude of the $O(1)$ term in (3.1). Choosing A sufficiently large that the right-hand side of (3.3) is strictly positive and observing that each term in the sum is at most $(A \ln x)/x$ establishes that there must be $\Omega(x/\log x)$ terms, and thus yields (3.2).

4. The Maximum Loss of Information

Our goal in this section is to determine the maximum possible loss of information in multipication. Our starting point is the formula (2.12). Since the average is at most the maximum, we have

$$H(X, Y \mid X \cdot Y) \leq \max_{1 \leq x \leq N} \max_{1 \leq y \leq N} \log_2 d(x \cdot y),$$

and since $\log_2 a$ is an increasing function of a, we obtain

$$H(X, Y \mid X \cdot Y) \leq \log_2 \left(\max_{1 \leq x \leq N} \max_{1 \leq y \leq N} d(x \cdot y) \right).$$

Using the fact that $d(x \cdot y) \leq d(x) \cdot d(y)$ as before, we obtain

$$H(X, Y \mid X \cdot Y) \leq 2 \log_2 \left(\max_{1 \leq n \leq N} d(n) \right). \qquad (4.1)$$

Wigert [W] was the first to show that

$$\log_2 \left(\max_{1 \leq n \leq N} d(n) \right) \sim \frac{\ln N}{\ln \ln N}, \qquad (4.2)$$

using the prime-number theorem. But Ramanujan [R] has shown that an estimate even more precise than (4.2) can be obtained using only the crude bounds

$$\pi(x) = \Theta \left(\frac{x}{\log x} \right) \qquad (4.3)$$

for the number $\pi(x)$ of primes not exceeding x obtained by Chebyshev [C]. Substituting (4.2) into (4.1) yields (1.4).

To show that this result is the best possible, we let X and Y be independent and uniformly distributed on the set $\mathcal{V}$ of the 2^k divisors of the product $v_k = p_1 \cdots p_k$ of the first k primes, where k is the largest integer such that

$$v_k \leq N. \qquad (4.4)$$

If we define $\vartheta(x)$ by

$$\vartheta(x) = \sum_{p \leq x} \ln p$$

(in which the sum is over primes p), then

$$v_k = \exp \vartheta(P_k),$$

so that (4.4) is equivalent to

$$\vartheta(p_k) \le \ln N.$$

The bounds

$$\vartheta(x) = \Theta(x)$$

are equivalent to the bounds (4.3) established by Chebyshev [C]. This implies that

$$p_k = \Theta(\log N),$$

so that (again using (4.3))

$$k = \Theta\left(\frac{\log N}{\log\log N}\right). \tag{4.5}$$

From (2.11), we have

$$H(X, Y \mid X \cdot Y) = \frac{1}{2^{2k}} \sum_{x\in\mathcal{V}} \sum_{y\in\mathcal{V}} H(X, Y \mid X \cdot Y = x \cdot y). \tag{4.6}$$

For $x, y \in \mathcal{V}$, let $u(x, y)$ denote the number of primes among $p_1, \ldots, p_k$ that divide one, but not both, of x and y. (This number is also the number of primes that divide the square-free part of $x \cdot y$, and thus it depends only on $x \cdot y$.) The random variable (X, Y), conditioned on $X \cdot Y = x \cdot y$, is uniformly distributed over the $2^{u(x,y)}$ pairs in the set

$$\mathcal{U} = \{(v, w) \in \mathcal{V} \times \mathcal{V} : v \cdot w = x \cdot y\},$$

so that

$$\begin{aligned} H(X, Y \mid X \cdot Y = x \cdot y) &= \log_2 \#\mathcal{U} \\ &= u(x, y). \end{aligned}$$

Thus (4.6) yields

$$H(X, Y \mid X \cdot Y) = \frac{1}{2^{2k}} \sum_{x\in\mathcal{V}} \sum_{y\in\mathcal{V}} u(x, y). \tag{4.7}$$

Since X and Y are each uniformly distributed on the 2^k divisors of v_k, the divisibility of each of X and Y by each of the primes $p_1, \ldots, p_k$ is probabilistically equivalent to the occurrences of heads among $2k$ independent flips of an unbiased coin. In particular, each of the primes $p_1, \ldots, p_k$ divides one, but not both, of X and Y with probability $1/2$. Thus the right-hand side of (4.7) is equal to $k/2$, and (4.5) yields

$$\begin{aligned} H(X, Y \mid X \cdot Y) &= k/2 \\ &= \Omega\left(\frac{\log N}{\log\log N}\right). \end{aligned}$$

This estimate shows that the result (1.4) is the best possible.

5. References

[A] H. Abelson and P. Andreae, “Information Transfer and Area-TimeTrade-Offs for VLSI Multiplication”, *Comm. ACM*, 23 (1980) 20–23.

[B] R. P. Brent and H. T. Kung, “The Area-Time Complexity of Binary Multiplication”, *Journal of the Association for Computing Machinery*, 28 (1981) 521–534; Corrigendum: 29 (1982) 904.

[C] P. L. Chebyshev (= Tchebichef), “Mémoire sur les nombres premiers”, *Journal de mathématiques pures et appliquées (1)*, 17 (1852) 366–390.

[D1] P. G. L. Dirichlet (= Lejeune-Dirichlet), “Sur l’usage des séries infinies dans la théorie des nombres”, *Journal für die reine und angewandte Mathematik”*, 18 (1838) 259–274.

[D2] P. G. L. Dirichlet (= Lejeune-Dirichlet), “Valeurs moyennes dans la théorie des nombres”, *Journal de mathématiques pures et appliquées (2)*, 1 (1956) 353–370.

[E1] P. Erdős, “On a New Method in Elementary Number Theory Which Leads to an Elementary Proof of the Prime Number Theorem”, *Proceedings of the National Academy of Science of the USA*, 35 (1949) 374–384.

[E2] P. Erdős, “Ob Odnom Asimptoticheskom Neravenstve v Teorii Chisel”, *Vestnik Leningradskogo Universiteta*, 13 (1960) 41–49.

[F] E. Fredkin and T. Toffoli, “Conservative Logic”, *International Journal of Theoretical Physics*, 21 (1982) 41–55.

[H1] J. Hadamard, “Sur la distribution des zéros de la fonction $\zeta(s)$ et ses conséquences arithmétiques”, *Bulletin de la Société Mathématiques de France”*, 24 (1896) 199–220.

[H2] G. H. Hardy and S. Ramanujan, “The Normal Number of Prime Factors of a Number n”, *Quarterly Journal of Mathematics*, 48 (1917) 76–92.

[M] F. Mertens, “Ein Beitrag zur analytischen Zahlentheorie”, *Journal für die reine und angewandte Mathematik”*, 78 (1874) 46–62.

[R] S. Ramanujan, “Highly Composite Numbers”, *Proceedings of the London Mathematical Society (2)*, 14 (1915) 347–409.

[S1] A. Selberg, “An Elementary Proof of the Prime-Number Theorem”, *Annals of Mathematics*, 50 (1949) 305–313.

[S2] A. Selberg, “An Elementary Proof of the Prime-Number Theorem for Arithmetic Progressions”, *Canadian Journal of Mathematics*, 2 (1950) 66–78.

[S3] C. E. Shannon, “A Mathematical Theory of Communication”, *Bell System Technical Journal*, 27 (1948) 379–423, 623–655.

[S4] H. N. Shapiro, “On Primes in Arithmetic Progression (II)”, *Annals of Mathematics*, 52 (1950) 231–243.

[V] Ch. de la Vallée Poussin, “Recherches analytiques sur la théorie des nombre premiers”, *Annales de la Société Scientifique de Bruxelles*, 20 (1896) 183–256, 281–397.

[W] S. Wigert, “Sur l’ordre de grandeur du nombre des diviseurs d’un entier”, *Arkiv för Matematik, Astronomi och Fysik*, 3, 18 (1907) 1–9.